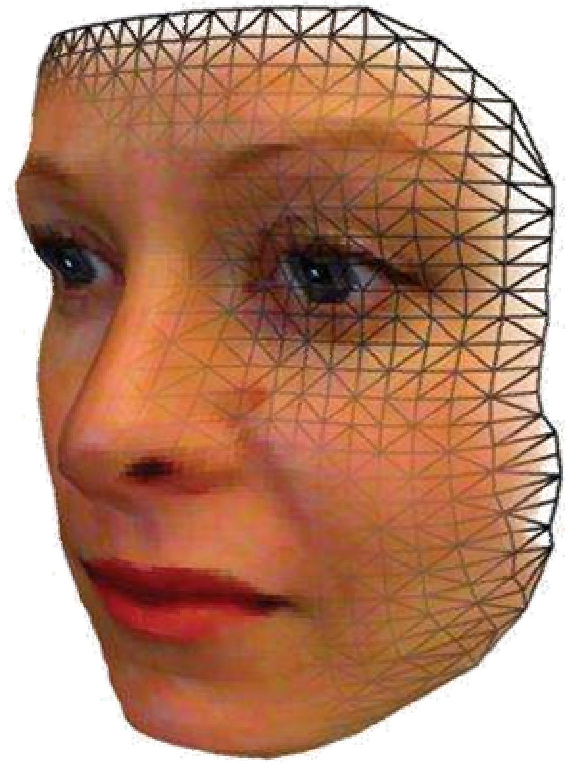
Human Face Recognition



PhD Proposal
Amirhosein Nabatchian

March 2008

Outline

- Research objective
- Introduction
- Applications of this research
- Challenges
- Literature review
- My approach
- Conclusions and future work

Research Objective

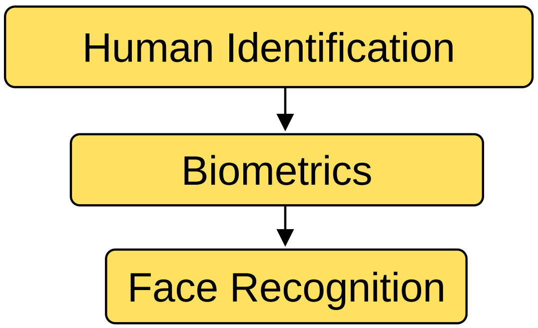
- Face recognition:

  Given still or video images of a scene, identify or verify one or more persons in the scene using a stored database of faces

- Areas of interest
  - Feature extraction and Classification fusion
  - Faces with non-uniform illumination
  - Faces with Occlusion
  - One sample per person problem

Introduction

- One of the most used biometrics
- Does not need participation
- There is feasible technology



Authentication vs. Identification

- Face Authentication/Verification (1:1 matching)

- Face Identification/Recognition (1:N matching)

Applications

Areas	Applications
Biometrics	Driver’s Licences, Entitlement Programs, Immigration, National ID, Passports, Voter Registration, Welfare Fraud
Information Security	Desktop Logon, Application Security, Database Security, File Encryption, Intranet Security, Internet Access, Medical Records, Secure Trading Terminals
Law Enforcement	Advanced Video Surveillance, CCTV Control, Portal Control, Post-Event Analysis, Shoplifting and Suspect Tracking and
Smart Cards	Stored Value Security, User Authentication
Access Control	Facility Access, Vehicular Access

Challenges

- automatically locate the face
- Identify similar faces (inter-class similarity)
- Accommodate intra-class variability due to:
  - head pose
  - illumination conditions
  - expressions
  - facial accessories
  - aging effects

What is Biometrics?

- The study of automated identification by use of physical or behavioral traits
  - Physical:
    Face, Fingerprint, Iris, Ear, Retina, Hands
  - Behavioral:
    Signature, Walking gait, Typing patterns
  - Both:
    Voice
- Biometric is used in all aspects of present day life for identifying individuals uniquely
  - ATM, car, house, work, computer
- A convenient and secure way of producing valid ID
- Can never forget it or lose it
- Growing field due to applicability to many aspects of everyday life

Essential Properties of a Biometric

- Universal
  Everyone should have the characteristic
- Uniqueness
  No two persons have the same characteristic
- Permanence
  Characteristic should be unchangeable
- Collectability
  Characteristic must be measurable

Psychophysics/Neuroscience Issues

- Face recognition is a dedicated process
- Both holistic and feature information are crucial for face perception and recognition
  - Human quickly focus on odd features
  - An inverted face is much harder to recognize
- Hair, face outline, eyes and mouth are important for perceiving faces
- Nose plays an insignificant role
  - But has the most features in profile face recognition
- Upper part of the face is much more important than the lower part
- pleasant faces, unpleasant faces, mid-range
- It's view point dependent
- Top lighting
- Movement (even if negated, inverted or thresholded)
- Facial expression is accomplished in parallel with face recognition

A Face Recognition System (FRS)

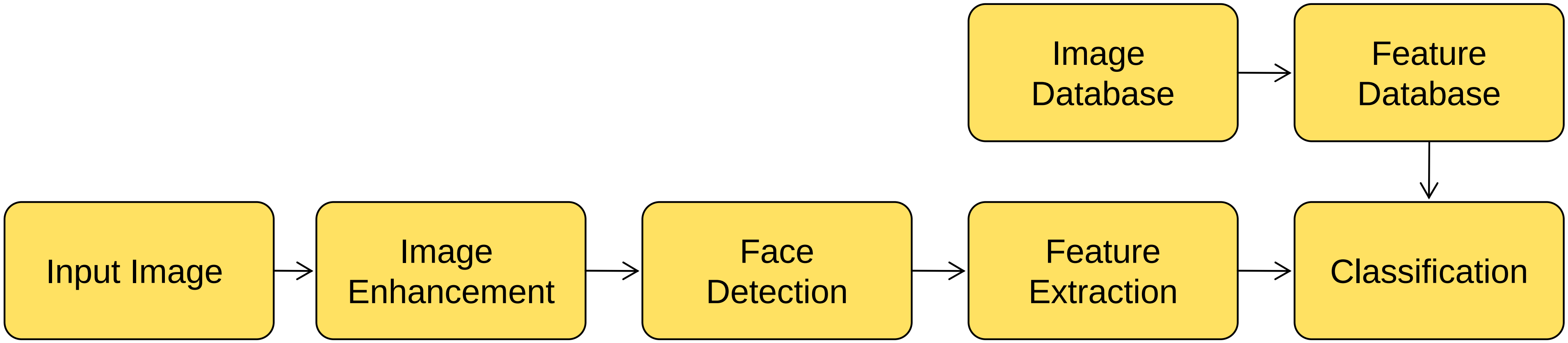


Image Enhancement

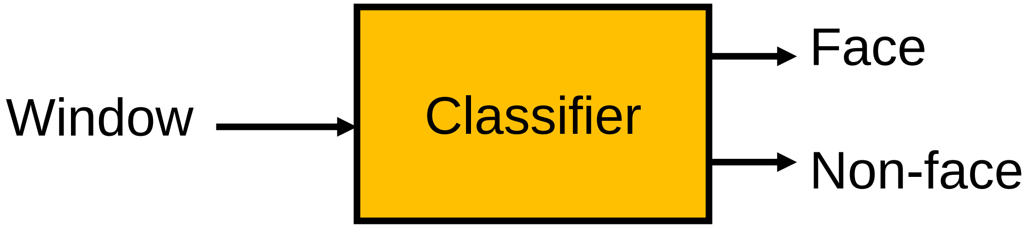
- Filtering & Noise removal
- Histogram equalization
- Color adjust

Face Detection/Segmentation

- Face template
- Deformable feature-based template
- Skin color
- Neural network
- Support Vector Machine (SVM)
- First order moment invariants

Face Detection

- Scan window over image
- Classify window as either:
  - Face
  - Non-face



An example

- Skin has a very small range of (intensity independent) colors, and little texture
  - Compute an intensity-independent color measure, check if color is in this range, check if there is little texture (median filter)
  - See this as a classifier - we can set up the tests by hand, or learn them.
  - get class conditional densities (histograms), priors from data (counting)

- Classifier is

  - if $p(\text{skin}|\boldsymbol{x}) > \theta$, classify as skin
  - if $p(\text{skin}|\boldsymbol{x}) < \theta$, classify as not skin
  - if $p(\text{skin}|\boldsymbol{x}) = \theta$, choose classes uniformly and at random

Feature Extraction

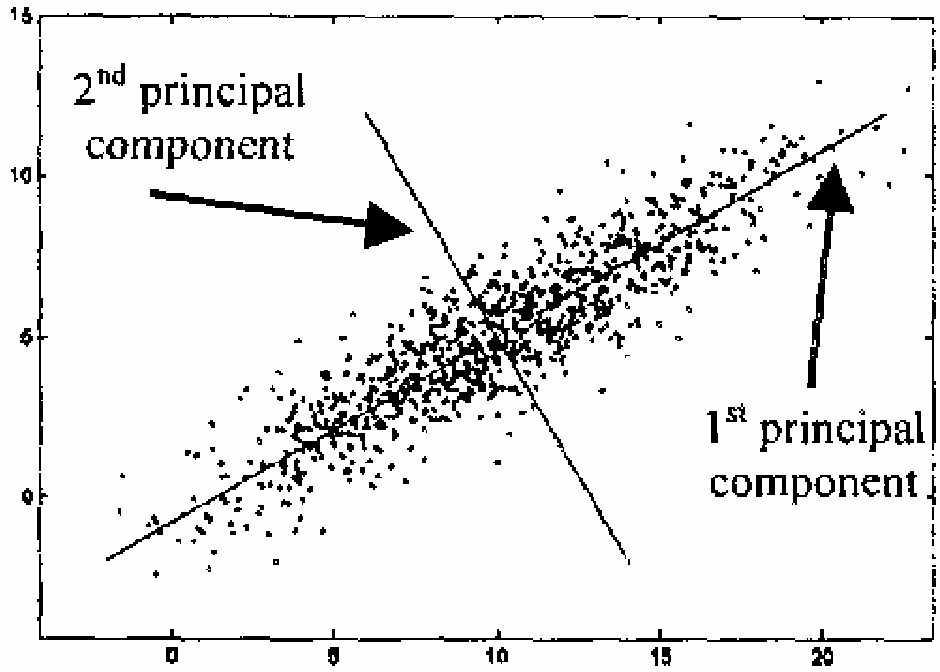
- Extraction of applicable features from the human face images is an important part of the recognition.
- There are two main approaches for this feature extraction problem
  - extracting geometrical and structural facial features like the shapes of the eyes, nose, mouth and the distance between these.
    - not affected by irrelevant information in the image, but sensitive to unpredictability of face appearance and environmental conditions.
  - the Holistic approach,
    - features from the whole image are extracted.
    - Since the features are global in the whole image, irrelevant elements like background and hair might affect the feature vectors and cause erroneous recognition results.

Feature Extraction

- Eigenfaces
- Generalized symmetry operator
  - Eyes, nose, mouth,
- Template-based eye detection
- Edge-based approach
- Gabor wavelet decomposition
- Moment Invariants

Eigenfaces

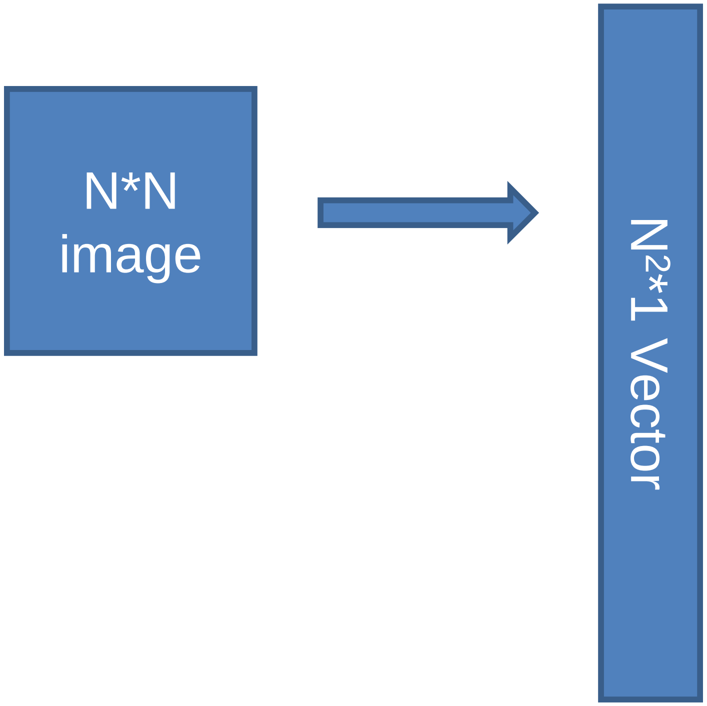
- Use Principle Component Analysis (PCA) to reduce the dimensionality
- Principal components are eigenvectors of covariance matrix



Eigenfaces

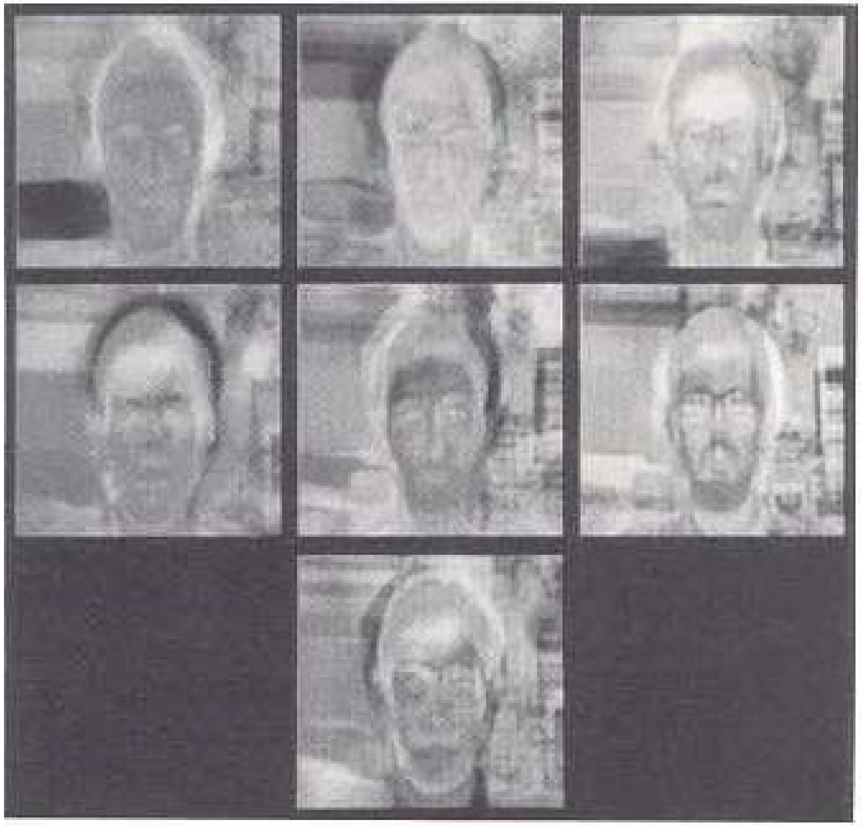
- Modeling
  1. Given a collection of *n* labeled training images,
  2. Compute mean image and covariance matrix. Subtract the mean.
  3. Compute k Eigenvectors (note that these are images) of covariance matrix corresponding to k largest Eigenvalues.
  4. Project the training images to the k-dimensional Eigenspace.
- Recognition
  1. Given a test image, project to Eigenspace.
  2. Perform classification to the projected training images.

Eigenfaces: Training Images



Eigenfaces

Mean Image



Basis Images

Difficulties with PCA

- Projection may suppress important detail
  - smallest variance directions may not be unimportant
- Method does not take discriminative task into account
  - typically, we wish to compute features that allow good discrimination
  - not the same as largest variance

PCA vs Fisher's Linear Discriminant

- Between-class scatter

$$S_B = \sum_{i=1}^{c} |\chi_i| (\mu_i - \mu)(\mu_i - \mu)^T$$

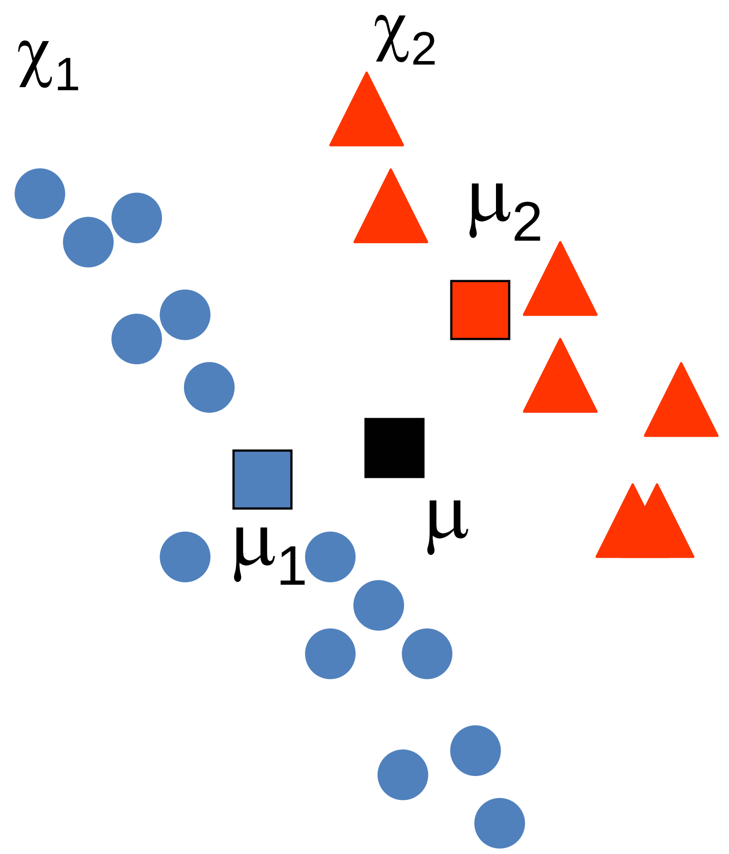
- Within-class scatter

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in \chi_i} (x_k - \mu_i)(\mu_k - \mu_i)^T$$

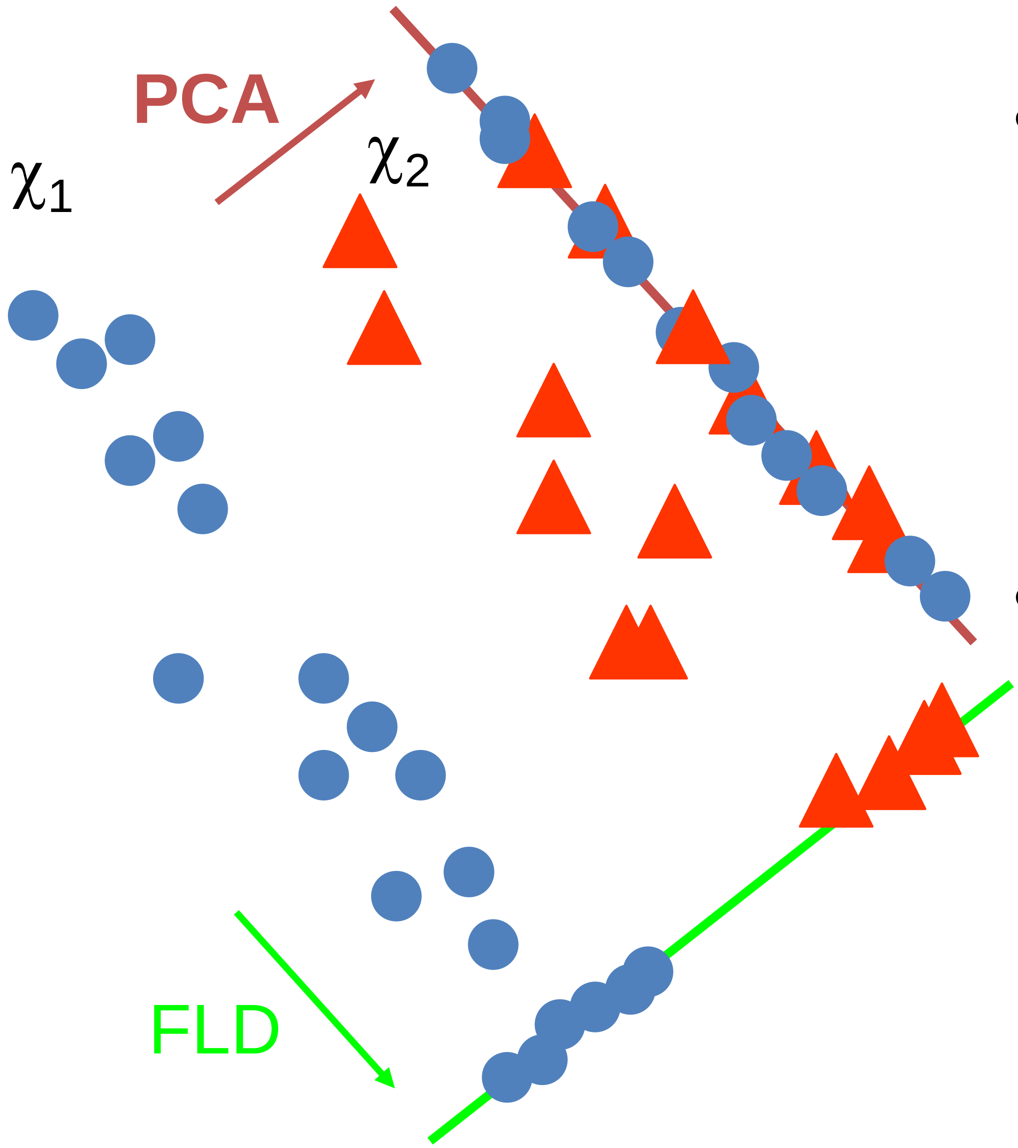
- Total scatter

$$S_T = \sum_{i=1}^{c} \sum_{x_k \in \chi_i} (x_k - \mu)(\mu_k - \mu)^T = S_B + S_W$$

- Where
  - c is the number of classes
  - μ_i is the mean of class χ_i
  - $|\chi_i|$ is number of samples of χ_i.



PCA & Fisher's Linear Discriminant



- PCA (Eigenfaces)

$$W_{PCA} = \arg\max_W \left| W^T S_T W \right|$$

Maximizes projected total scatter

- Fisher's Linear Discriminant

$$W_{fld} = \arg\max_W \frac{\left| W^T S_B W \right|}{\left| W^T S_W W \right|}$$

Maximizes ratio of projected between-class to projected within-class scatter

Moment Invariants

- Moment features are invariant under scaling, translation rotation and reflection
- Moment Invariants have been proven to be useful in pattern recognition applications due to their sensitivity to the pattern features

Moment invariants

- In 1961 Hu introduced the first set of algebraic moment invariants
- Simply when calculating an image's center we are using first order Regular moment invariants (RMI).
- Bamieh and De Figueiredo derived another set of moment invariants (BMI) which have small feature vectors that are computationally efficient.
- Zernike and pseudo Zernike orthogonal polynomials are the basis of the Zernike and pseudo Zernike moment invariants (ZMI) and (PZMI).
- Teague-Zernike Moment invariants (TZMI) were proposed by Teauge as a method for calculating ZMIs.
- Zernike and pseudo Zernike moment invariants have been normalized to generate NZMI and NPMZIs. The normalized forms have yielded better results in simple recognition applications

Regular Moment Invariants

$$M_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dxdy$$

$$\overline{x} = \frac{M_{10}}{M_{00}}, \qquad \overline{y} = \frac{M_{01}}{M_{00}}$$

$$M_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \overline{x})^p (y - \overline{y})^q f(x, y) dxdy$$

$$\mu_{pq} = \frac{1}{\left[M_{20} + M_{02}\right]^{\gamma}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \overline{x})^p (y - \overline{y})^q f(x, y) dxdy$$

Regular Moment Invariants

$$(RMI)_{jk} = \sum_{r=0}^{j}\sum_{s=0}^{k}(-1)^{k-s}\binom{j}{r}\binom{k}{s}(\cos\theta)^{j-r+s} \times (\cos\theta)^{k+r-s}\,\mu_{(j+k-r-s),(r+s)}$$

$$\theta = \frac{1}{2}\tan^{-1}\left[\frac{2M_{11}}{M_{20}-M_{02}}\right]$$

Hu Moment Invariants

$$I_{p-r,r} = \sum_{l=0}^{r} (-j)^l \begin{bmatrix} p-2l \\ l \end{bmatrix} \sum_{k=0}^{r} \begin{bmatrix} r \\ k \end{bmatrix} \mu_{p-2k-l,2k+l}$$

$$p-2r > 0, \quad j = \sqrt{-1}$$

$$(HMI)_r = \left| I_{p-r,r} \right|^2, r = 1, 2, 3, \ldots, p-2r$$

Bamieh Moment Invariants

$$M^{ijk} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^i x^j x^k \ldots f\left(x^1, x^2\right) dx^1 dx^2$$

2nd order	$(BMI)_1 = \mu_{20}\mu_{02} - \mu_{11}{}^2$
3rd order	$(BMI)_2 = \left(\mu_{30}\mu_{03} - \mu_{12}\mu_{21}\right)^2 - 4\left(\mu_{03}\mu_{12} - \mu_{21}{}^2\right)\left(\mu_{21}\mu_{30} - \mu_{12}{}^2\right)$
4th order	$(BMI)_3 = \mu_{40}\mu_{04} - 4\mu_{13}\mu_{31} + 3\mu_{22}{}^2$ $(BMI)_4 = \mu_{40}\mu_{22}\mu_{04} - 2\mu_{13}\mu_{22}\mu_{31} - \mu_{13}{}^2\mu_{40} - \mu_{04}\mu_{31}{}^2 - \mu_{22}{}^3$

Zernike Moment Invariants

$$A_{nL} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^{\infty} f\left(r\cos\varphi, r\sin\varphi\right) R_{nL}\left(r\right) \times e^{-jL\varphi} r\,dr\,d\varphi$$

$$R_{nL}\left(r\right) = \sum_{k=L}^{n} B_{nLk} r^k$$

$$B_{nLk} = \frac{\left(-1\right)^{\frac{n-k}{2}} \times \left(\frac{n+k}{2}\right)!}{\left(\frac{n-k}{2}\right)! \times \left(\frac{L+k}{2}\right)! \times \left(\frac{k-L}{2}\right)!}$$

Pseudo Zernike Moment Invariants

$$
\begin{aligned}
PZMI_{nm} = &\\
& \frac{n+1}{\pi} \sum_{\substack{s=0 \\ (n-m-s)even}}^{n-|m|} D_{n,|m|,s} \\
& \times \sum_{a=0}^{k} \sum_{b=0}^{m} \binom{k}{a} \binom{m}{b} (-j)^b \, CM_{2k+m-2a-b,2a+b} \\
& + \frac{n+1}{\pi} \sum_{\substack{s=0 \\ (n-m-s)odd}}^{n-|m|} D_{n,|m|,s} \\
& \times \sum_{a=0}^{d} \sum_{b=0}^{m} \binom{d}{a} \binom{m}{b} (-j)^b \, RM_{2d+m-2a-b,2a+b}
\end{aligned}
$$

$$
RM_{pq} = \frac{\sum_x \sum_y f(x,y) \hat{x}^p \hat{y}^q \sqrt{\hat{x}^2 + \hat{y}^2}}{M_{00}^{(p+q+2)/2}}
$$

$$
CM_{pq} = \frac{\mu_{pq}}{M_{00}^{(p+q+2)/2}}
$$

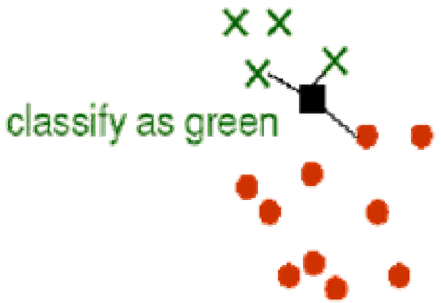
$$
D_{n,|m|,s} = \frac{(-1)^s \times (2n+1-s)!}{s! \times (n-|m|-s)! \times (n-|m|-s+1)!}
$$

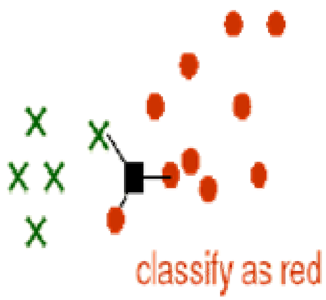
Classification

- *k*-NN
- Neural Network
- SVM
- RBF
- Fuzzy Logic

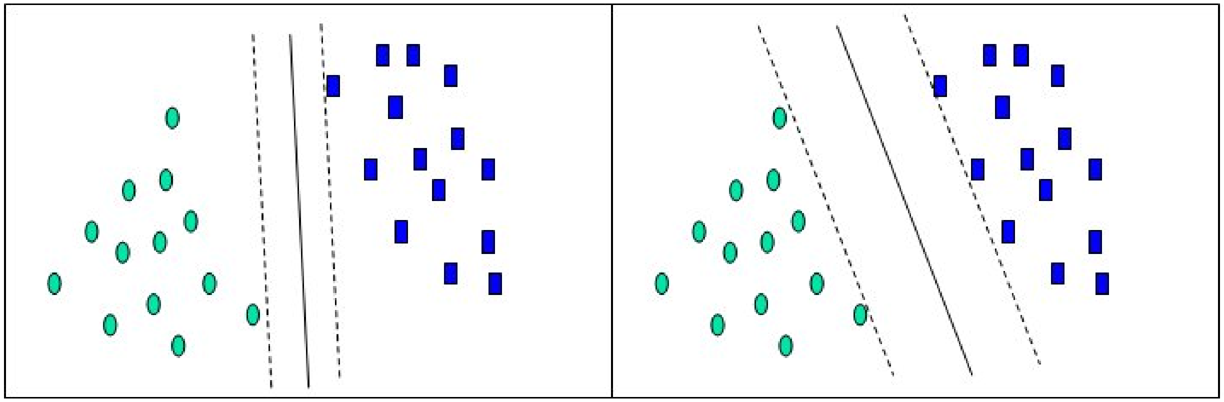
k-NN

- K-nearest neighbors:
  - Given x, take vote among its k nearest neighbors
- Advantages:
  - Training is very fast
  - Learns complex target functions
  - Does not loose information
- Disadvantages:
  - Slow at query time
  - Easily fooled by irrelevant attributes





Support Vector Machines



Small Margin Large Margin

• When not linearly separable, transform to a higher order space, separate linearly, and transform to the output space

The Database

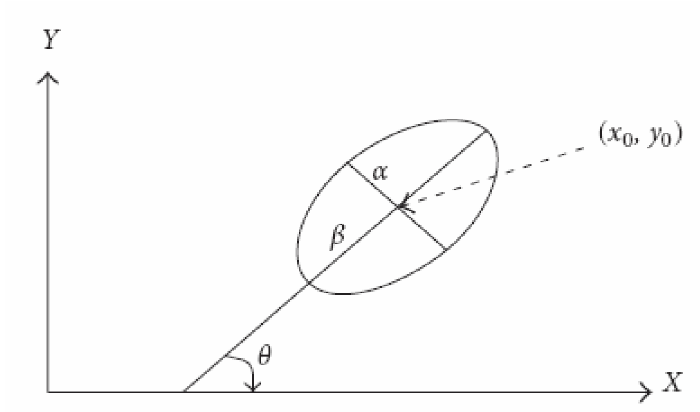
- The AT&T face image database (formerly known as the ORL database)
- A set of face images taken between April 1992 and April 1994 from 40 people
- 10 different images of each person taken at different time, illuminations, and facial expressions like open or closed eyes, smiling or not smiling and facial details like glasses.
- dark homogeneous background, the subject is in an upright, frontal position.
- The size of each image is 92×112 pixels, with 256 grey levels per pixel

The Database

LOCATING THE FACE AND CREATING SUBIMAGES

- we use second order moment invariants

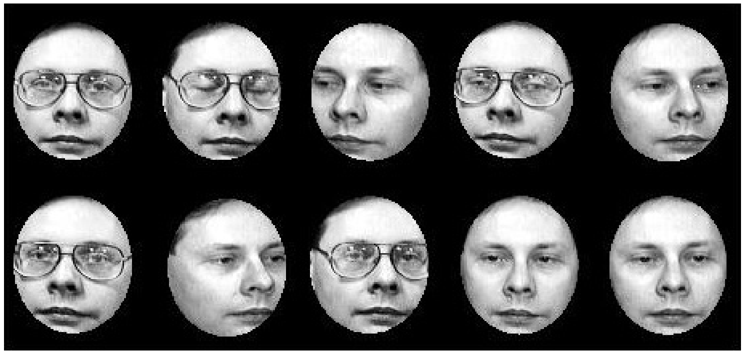
$$\alpha = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{\mathrm{I}_{\max}^3}{\mathrm{I}_{\min}}\right)^{\frac{1}{8}}, \beta = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{\mathrm{I}_{\min}^3}{\mathrm{I}_{\max}}\right)^{\frac{1}{8}}$$

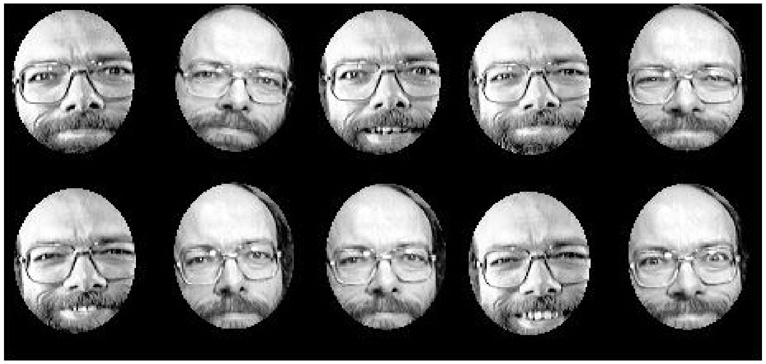


$$\mathrm{I}_{\min} = \sum_x \sum_y \left[(x - x_0)\cos\theta - (y - y_0)\sin\theta\right]^2$$

$$\mathrm{I}_{\max} = \sum_x \sum_y \left[(x - x_0)\sin\theta - (y - y_0)\cos\theta\right]^2$$

- result of face localization on images of the database





Feature Vector and Feature Matrix

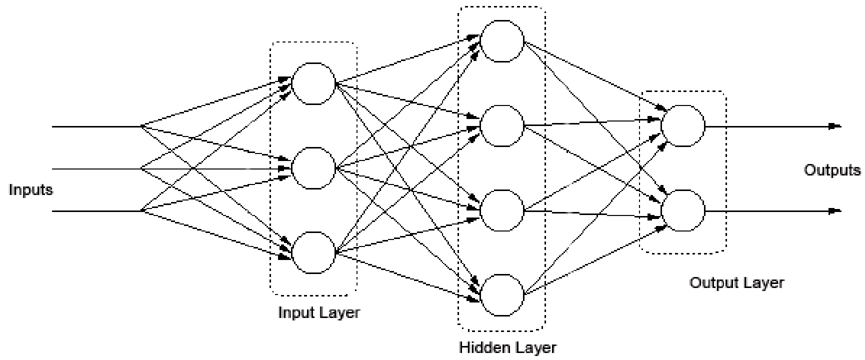
$$FV_{MI,n} = \left[MI_1, MI_2, MI_3, ..., MI_n \right]$$

$$FV_{BMI} = \left[BMI_1, BMI_2, BMI_3, BMI_4 \right]$$

$$WFV_{MI,n} = \frac{FV_{MI,n} - mean\left(FV_{MI,n}\right)}{\mathrm{var}\left(FV_{MI,n}\right)}$$

Calculating this vector for all the images in our training database will result in a $M \times n$ feature space matrix where M is the number of images in the training database and n is the number of the moments that we are going to use for each type of moment invariants.

- Artificial neural networks (ANN) have been widely used as the classifier in many face recognition systems
- We have used a three layer perceptron neural network.
- The number of neurons in input layer is equal to the size of related feature vector for each experiment
- The number of output neurons is equal to the number of classes which is 40



Recognition Results

classification parameters for different moment invariants

Method	Classification accuracy	Number of features	order	Number of neurons in input layer	Number of neurons in hidden layer	R	Number of epochs
PZMI	94 %	120	10	120	108	0.9	1000
NZMI	91 %	34	7	34	68	2.0	6000
NPZMI (10)	90.5 %	118	10	118	106	0.9	1000
ZMI	87 %	76	11	76	152	2.0	4000
NPZMI (7)	86 %	61	7	61	122	2.0	2000
RMI	85.5 %	75	11	75	150	2.0	2000
TZMI	67.5 %	32	7	32	96	3.0	10000
HMI	61 %	32	7	32	96	3.0	6000
BMI	22 %	4	4	4	12	3.0	4000

Recognition Speed

recognition time for different methods

Moment Invariant	Recognition time (s)	order	Number of Features
BMI	0.15	4	4
HMI	0.62	7	32
RMI	4.53	7	33
PZMI	10.75	7	64 (36 Re+28 Im)
ZMI	124	7	33 (18 Re+15 Im)

Conclusion

- Performance of different moment invariants for feature extraction in face recognition was studied
- The moment invariants used in this study were HMI, BMI, RMI, ZMI, TZMI, PZMI, NZMI and NPZMI.
- A three layer perceptron neural network was used as the classifier in the recognition system.
- The performance of the system can be optimized by using proper value for ρ and also by choosing proper orders of the moment invariants.
- High order PZMIs contain useful information, therefore lead to the best results.
- BMIs were the fastest to be computed but had very poor recognition accuracy for face images
- calculating ZMIs was the most time consuming.
- The best recognition rate of 95% was achieved with PZMI of order 1-7.

Future work

- Feature extractor and classifier fusion
- Non-uniformed illumination
- Occlusion
- one sample problem

References

- [1] Heisele, B., Ho, P., Wu, J., Poggio, T., "Face recognition: component-based versus global approaches", *Computer Vision and Image Understanding*, Vol. 91 No.1/2, pp.6-21.2003
- [2] Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A. (2003), "Face recognition: a literature survey", *ACM Computing Surveys*, Vol. 35 No.4, pp.339-458.
- [3] J. Haddadnia, M.,Ahmadi, K.,Faez.: An efficient feature extraction method with pseudo-Zernike moment in RBF neural network-based human face recognition system. EURASIP Journal on Applied Signal Processing (2003), 890–901.
- [4] E. Hjelmas and B. K. Low, “Face detection: a survey,” *Computer Vision and Image Understanding*, vol. 83, no. 3, pp. 236– 274, 2001.
- [5] M. Hu, visual Pattern recognition by Moment Invariants, IRE Trans. Inf. Theory 8, 179-187, 1962
- [6] S. Reddi, Radial and Angular Moment Invariants for image Identification, IEEE Trans. Pattern Analysis and Machine intelligence, 3,240-242, 1981
- [7] R.Bamieh, R. De Figueiredo, A General Moment Invariants/Attributed-graph method for the three dimensional object recognition from a single image, IEEE Journal of Robotics Automation 2, 31-41,1986
- [8] C. Teh, R. T. Chin, On Image Analysis by the Methods of Moments, IEEE Transactions on Pattern Analysis and Machine Intelligence, v.10 n.4, p.496-513, July 1988
- [9] M. Teague, "Image analysis via the general theory of moments,"J. Opt. Soc. Amer., vol. 70, no. 8, pp. 920-930, Aug. 1980.

References

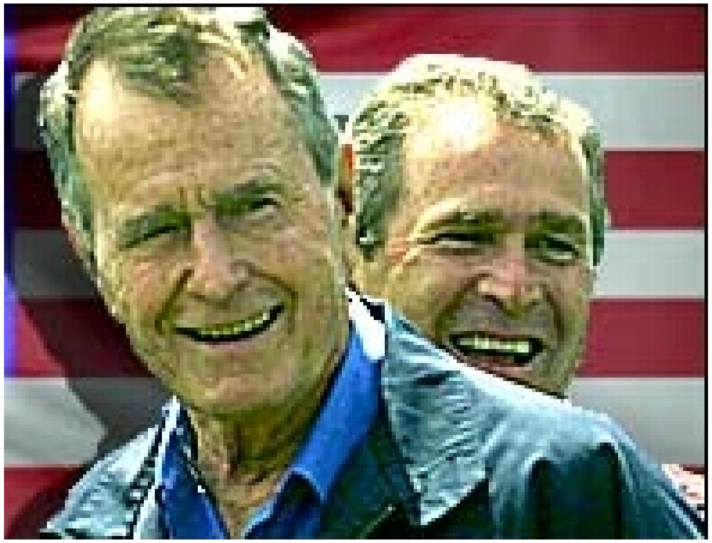
- [10] S. O. Belkasim , M. Shridhar , M. Ahmadi, Pattern recognition with moment invariants: a comparative study and new results, Pattern Recognition, v.24 n.12, p.1117-1138, Dec. 1991
- [11] http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
- [12] J. Wang and T. Tan, "A new face detection method based on shape information," *Pattern Recognition Letter*, vol. 21, no. 6-7, pp. 463–471, 2000.
- [13] J. Haddadnia and K. Faez, "Human face recognition using radial basis function neural network," in *Proc. 3rd International Conference On Human and Computer (HC '00)*, pp. 137–142, Aizu, Japan, September 2000.
- [14] J. Haddadnia, M. Ahmadi, and K. Faez, "A hybrid learning RBF neural network for human face recognition with pseudo Zernike moment invariant," in *IEEE International Joint Conference On Neural Network (IJCNN '02)*, pp. 11–16, Honolulu, Hawaii, USA, May 2002.
- [15] S. Gutta, J. R. J. Huang, P. Jonathon, and H. Wechsler, "Mixture of experts for classification of gender, ethnic origin, and pose of human faces," *IEEE Transactions on Neural Networks*, vol. 11, no. 4, pp. 948–960, 2000.
- [16] S. Z. Li and J. Lu, "Face recognition using the nearest feature line method," *IEEE Transactions on Neural Networks*, vol. 10, no. 2, pp. 439–443, 1999.
- [17] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a convolutional neural network approach," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, 1997.

Thanks for your kind attention

Inter-Class Similarity

- Different persons may have very similar appearance

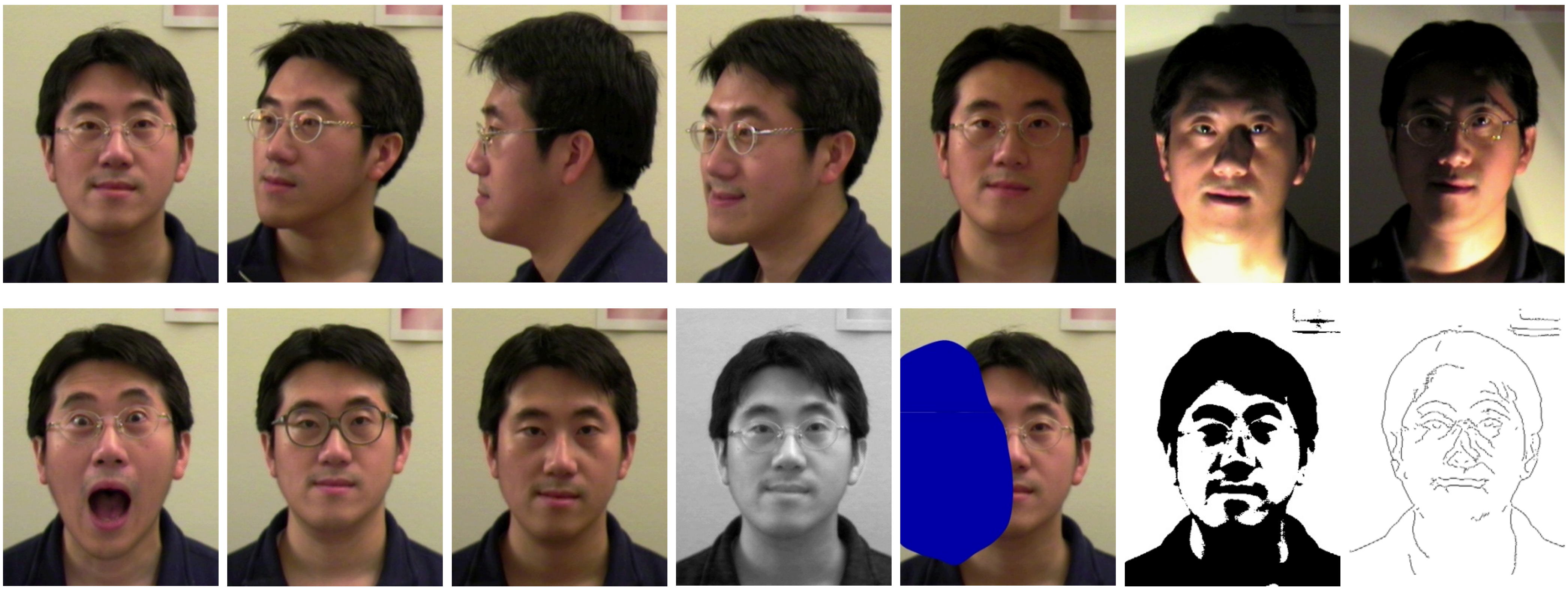
Twins



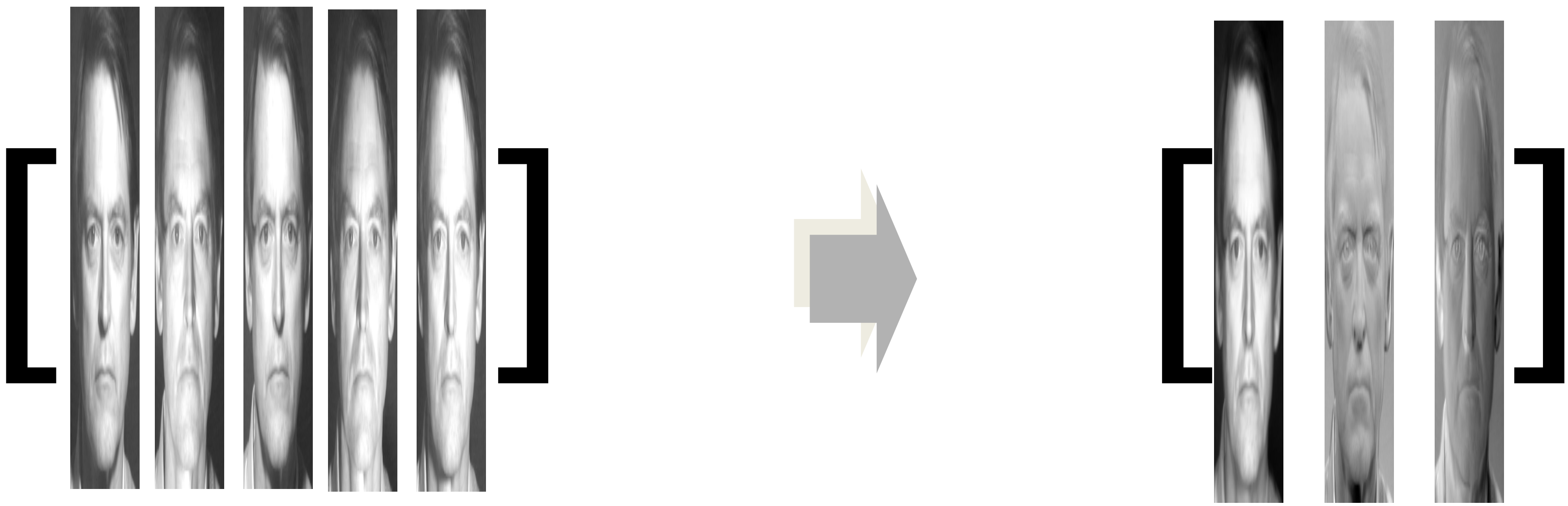
Father and son

Intra-Class Variability

- Faces with intra-subject variations in pose, illumination, expression, accessories, color, occlusions, and brightness



How to construct Eigenspace?



$[\ \mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3\ \mathbf{x}_4\ \mathbf{x}_5\ ]$

$\mathbf{W}$

Construct data matrix by stacking vectorized images and then apply Singular Value Decomposition (SVD)

Fisherfaces: Class specific linear projection

- An n-pixel image $\boldsymbol{x} \in \mathbf{R}^n$ can be projected to a low-dimensional feature space $\boldsymbol{y} \in \mathbf{R}^m$ by

$$\boldsymbol{y} = W\boldsymbol{x}$$

where W is an n by m matrix.

- Recognition is performed using nearest neighbor in $\mathbf{R}^m$.

- How do we choose a good W?